

USING THE SHERMAN-MORRISON-WOODBURY FORMULA TO SOLVE THE SYSTEM OF LINEAR EQUATIONS FROM THE STANDARD MULTIPLE SHOOTING METHOD FOR A LINEAR TWO POINT BOUNDARY-VALUE PROBLEM IS A BAD IDEA

IVO HEDTKE

ABSTRACT. We use the standard multiple shooting method to solve a linear two point boundary-value problem. To ensure that the solution obtained by combining the partial solutions is continuous and satisfies the boundary conditions, we have to solve a system of linear equations. Our idea is to first solve a bidiagonal system related to the original system of linear equations, and then update it with the Sherman-Morrison-Woodbury formula. We study the feasibility, the numerical stability and the running time of this method. The results are: The method described above has the same stability problems like the well known Condensing method. The running time analysis shows that the new method is slower than the Condensing method. Therefore we recommend not to use the method described in this article.

1. INTRODUCTION

We solve the linear two point boundary-value problem

$$\begin{aligned}\mathcal{L}\mathbf{x}(t) &:= \dot{\mathbf{x}}(t) - \mathbf{A}(t)\mathbf{x}(t) = \mathbf{r}(t), & t \in [a, b] \\ \mathcal{B}\mathbf{x}(t) &:= \mathbf{B}_a\mathbf{x}(a) + \mathbf{B}_b\mathbf{x}(b) = \boldsymbol{\beta}\end{aligned}$$

with the standard multiple shooting method, where $\mathbf{x}(t)$, $\mathbf{r}(t): [a, b] \rightarrow \mathbb{R}^n$, $\boldsymbol{\beta} \in \mathbb{R}^n$, $\mathbf{A}(t): [a, b] \rightarrow \mathbb{R}^{n \times n}$ and $\mathbf{B}_a, \mathbf{B}_b \in \mathbb{R}^{n \times n}$. We divide the interval $[a, b]$ with the shooting points

$$a = \tau_0 < \tau_1 < \dots < \tau_{m-1} < \tau_m = b$$

into m segments $[\tau_j, \tau_{j+1}]$. We use the principle of superposition on each segment to find the solution

$$\mathbf{x}_j(t) = \mathbf{X}(t; \tau_j)\mathbf{c}_j + \mathbf{v}(t; \tau_j),$$

where \mathbf{c}_j is a constant vector. $\mathbf{X}(t; \tau_j)$ is a fundamental system which fulfills the IVP

$$\mathcal{L}\mathbf{X}(t; \tau_j) = \mathbf{0}, \quad t \in [\tau_j, \tau_{j+1}], \quad \mathbf{X}(\tau_j; \tau_j) = \mathbf{I}.$$

$\mathbf{v}(t; \tau_j)$ is an inhomogeneous solution of the ODE and fulfills

$$\mathcal{L}\mathbf{v}(t; \tau_j) = \mathbf{r}(t), \quad t \in [\tau_j, \tau_{j+1}], \quad \mathbf{v}(\tau_j; \tau_j) = \mathbf{0}.$$

The problem now consists in determining the vectors \mathbf{c}_j in such a way, that

- (1) the function $\mathbf{x}(t)$ pieced together by the $\mathbf{x}_j(t)$ is continuous and
- (2) satisfies the boundary conditions.

We define $\mathbf{X}_j := \mathbf{X}(\tau_{j+1}; \tau_j)$ and $\mathbf{v}_j := \mathbf{v}(\tau_{j+1}; \tau_j)$. To satisfy the boundary conditions we focus on $\mathcal{B}\mathbf{x}(t) = \boldsymbol{\beta}$:

$$(1) \quad \mathbf{B}_a\mathbf{c}_0 + \mathbf{B}_b\mathbf{X}_{m-1}\mathbf{c}_{m-1} = \boldsymbol{\beta} - \mathbf{B}_b\mathbf{v}_{m-1}.$$

To ensure that $\mathbf{x}(t)$ is a continuous function we need

$$\mathbf{x}_{k-1}(\tau_k) = \mathbf{x}_k(\tau_k), \quad k = 1, \dots, m-1,$$

2000 *Mathematics Subject Classification.* 34B05, 15A06.

Key words and phrases. Sherman-Morrison-Woodbury formula, standard multiple shooting method, linear two point boundary-value problem, Condensing method.

The author was supported by the Studienstiftung des Deutschen Volkes.

which yields to the conditions

$$(2) \quad \mathbf{c}_k - \mathbf{X}_{k-1} \mathbf{c}_{k-1} = \mathbf{v}_{k-1}, \quad k = 1, \dots, m-1.$$

Now we collect equation (1) and the $m-1$ equations (2) in the following system of linear equations:

$$(3) \quad \mathbf{M}\mathbf{c} = \mathbf{q},$$

where we define $\mathbf{Y}_j := -\mathbf{X}_j$ and

$$\mathbf{M} := \begin{bmatrix} \mathbf{Y}_0 & \mathbf{I} & & \\ & \mathbf{Y}_1 & \mathbf{I} & \\ & & \ddots & \ddots \\ & & & \mathbf{Y}_{m-2} & \mathbf{I} \\ \mathbf{B}_a & & & & \mathbf{B}_b \mathbf{X}_{m-1} \end{bmatrix}, \quad \mathbf{c} := \begin{pmatrix} \mathbf{c}_0 \\ \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_{m-1} \end{pmatrix}, \quad \mathbf{q} := \begin{pmatrix} \mathbf{v}_0 \\ \vdots \\ \mathbf{v}_{m-2} \\ \beta - \mathbf{B}_b \mathbf{v}_{m-1} \end{pmatrix}.$$

Note that $\mathbf{c}, \mathbf{q} \in \mathbb{R}^{mn}$ and $\mathbf{M} \in \mathbb{R}^{mn \times mn}$. It is known that \mathbf{M} is regular if we assume that the BVP has an unique solution. In this case

$$(4) \quad \mathbf{N} := \mathbf{B}_a + \mathbf{B}_b \mathbf{X}(b; a)$$

is regular, too. (see [2, Satz 8.1 (Theorem 8.1)])

2. THE AIM OF THIS WORK

There exists the well known method *Condensing* to solve the system (3) (see Section 6). Because of the special structure of \mathbf{M} it is pretty obvious to try to find the solution in the following way: First solve the bidiagonal system from (6) and then update the solution with the Sherman-Morrison-Woodbury formula. In this paper we study the feasibility, the numerical stability and the running time of this method.

3. THE SHERMAN-MORRISON-WOODBURY FORMULA

Let \mathbf{A} be a regular $\ell \times \ell$ matrix and \mathbf{U} and \mathbf{V} be two $\ell \times p$ matrices. If $\mathbf{I}_p + \mathbf{V}^\top \mathbf{A}^{-1} \mathbf{U}$ is regular, then

$$(5) \quad (\mathbf{A} + \mathbf{U} \mathbf{V}^\top)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{U} (\mathbf{I}_p + \mathbf{V}^\top \mathbf{A}^{-1} \mathbf{U})^{-1} \mathbf{V}^\top \mathbf{A}^{-1}.$$

holds.

4. IS IT POSSIBLE TO USE THE SHERMAN-MORRISON-WOODBURY FORMULA TO SOLVE $\mathbf{M}\mathbf{c} = \mathbf{q}$?

First, we have to split \mathbf{M} into two matrices $\mathbf{M} = \mathcal{M} + \mathcal{U}$, where \mathcal{U} can be written in the form $\mathcal{U} = \mathbf{U} \mathbf{V}^\top$ with $\mathbf{U}, \mathbf{V} \in \mathbb{R}^{mn \times n}$. For this we define

$$\mathbf{U} = [\mathbf{0}, \dots, \mathbf{0}, \mathbf{B}_a^\top]^\top \quad \text{and} \quad \mathbf{V}^\top = [\mathbf{I}_n, \mathbf{0}, \dots, -\mathbf{L}],$$

where $\mathbf{L} := \mathbf{X}_0^{-1} \cdots \mathbf{X}_{m-2}^{-1}$. Therefore we have

$$\mathcal{U} = \mathbf{U} \mathbf{V}^\top = \begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} \\ & \ddots & \\ \mathbf{B}_a & & -\mathbf{B}_a \mathbf{L} \end{bmatrix},$$

and

$$(6) \quad \mathcal{M} = \mathbf{M} - \mathcal{U} = \begin{bmatrix} \mathbf{Y}_0 & \mathbf{I} & & \\ & \mathbf{Y}_1 & \mathbf{I} & \\ & & \ddots & \ddots & \mathbf{I} \\ & & & \mathbf{Y}_{m-2} & \mathbf{I} \\ \mathbf{B}_a & & & & \mathcal{B} \end{bmatrix},$$

where $\mathcal{B} := \mathbf{B}_b \mathbf{X}_{m-1} + \mathbf{B}_a \mathbf{L}$.

Now we have to check that \mathcal{M} is regular. Because of

$$\det \mathcal{M} = \det \mathcal{B} \prod_{j=0}^{m-2} \det \mathbf{Y}_j,$$

it follows that $\det \mathcal{M} \neq 0$ iff $\det \mathcal{B} \neq 0$, because the \mathbf{Y}_j are fundamental systems. But $\mathcal{B} = \mathbf{N}\mathbf{L}$ and \mathbf{N} and \mathbf{L} are both regular. This follows from

$$(7) \quad \mathbf{X}(b; a) = \prod_{j=1}^m \mathbf{X}_{m-j}.$$

This shows that \mathcal{M} is regular.

Finally we have to check that $\mathbf{I}_n + \mathbf{V}^\top \mathcal{M}^{-1} \mathbf{U}$ is regular. First we need an auxiliary result:

Lemma. *Given m regular $n \times n$ matrices \mathbf{D}_i . Then, the matrix*

$$\Delta := \begin{bmatrix} \mathbf{D}_0 & \mathbf{I}_n \\ & \mathbf{D}_1 & \mathbf{I}_n \\ & & \ddots & \ddots \\ & & & \mathbf{D}_{m-2} & \mathbf{I}_n \\ & & & & \mathbf{D}_{m-1} \end{bmatrix}$$

is regular and

$$\Delta^{-1} = \begin{bmatrix} \mathbf{D}_0^{-1} & -(\mathbf{D}_1 \mathbf{D}_0)^{-1} & (\mathbf{D}_2 \mathbf{D}_1 \mathbf{D}_0)^{-1} & \dots & (-1)^{m-1} (\mathbf{D}_{m-1} \cdots \mathbf{D}_0)^{-1} \\ & \mathbf{D}_1^{-1} & -(\mathbf{D}_2 \mathbf{D}_1)^{-1} & & \\ & & \ddots & & \\ & & & \mathbf{D}_{m-2}^{-1} & -(\mathbf{D}_{m-1} \mathbf{D}_{m-2})^{-1} \\ & & & & \mathbf{D}_{m-1}^{-1} \end{bmatrix}$$

holds. \square

Proof. It holds $\det \Delta = \prod_{j=0}^{m-1} \det \mathbf{D}_j \neq 0$. $\Delta \Delta^{-1} = \mathbf{I}_{mn}$ and $\Delta^{-1} \Delta = \mathbf{I}_{mn}$ can easily be verified. \square

Now we go back to the matrix $\mathbf{I}_n + \mathbf{V}^\top \mathcal{M}^{-1} \mathbf{U}$. With \mathcal{M}_j^{-1} we denote the j th column of \mathcal{M}^{-1} and we write \mathcal{M}_{ij}^{-1} for the $n \times n$ sub-matrix in the i th row and j th column of \mathcal{M}^{-1} . With the lemma above and the new notation we get

$$\begin{aligned} \mathbf{V}^\top \mathcal{M}^{-1} \mathbf{U} &= [\mathbf{I}_n, \mathbf{0}, \dots, \mathbf{0}, -\mathbf{L}] [\mathcal{M}_1^{-1} \mid \dots \mid \mathcal{M}_m^{-1}] [\mathbf{0}, \dots, \mathbf{0}, \mathbf{B}_a^\top]^\top \\ &= [\mathcal{M}_{11}^{-1} - \mathbf{L} \mathcal{M}_{m1}^{-1} \mid \dots \mid \mathcal{M}_{1m}^{-1} - \mathbf{L} \mathcal{M}_{mm}^{-1}] [\mathbf{0}, \dots, \mathbf{0}, \mathbf{B}_a^\top]^\top \\ &= \mathcal{M}_{1m}^{-1} \mathbf{B}_a - \mathbf{L} \mathcal{M}_{mm}^{-1} \mathbf{B}_a. \end{aligned}$$

With the special structure of \mathcal{M}^{-1} we can calculate the two sub-matrices \mathcal{M}_{1m}^{-1} and \mathcal{M}_{mm}^{-1} very easy: $\mathcal{M}_{mm}^{-1} = \mathcal{B}^{-1}$ and

$$\begin{aligned} \mathcal{M}_{1m}^{-1} &= (-1)^{m-1} \left(\mathcal{B} \prod_{j=2}^m \mathbf{Y}_{m-j} \right)^{-1} = (-1)^{m-2} \mathbf{Y}_0^{-1} \cdots \mathbf{Y}_{m-2}^{-1} \mathcal{B}^{-1} \\ &= \mathbf{X}_0^{-1} \cdots \mathbf{X}_{m-2}^{-1} \mathcal{B}^{-1} = \mathbf{L} \mathcal{B}^{-1}. \end{aligned}$$

Now it follows that

$$\mathbf{V}^\top \mathcal{M}^{-1} \mathbf{U} = \mathcal{M}_{1m}^{-1} \mathbf{B}_a - \mathbf{L} \mathcal{M}_{mm}^{-1} \mathbf{B}_a = \mathbf{L} \mathcal{B}^{-1} \mathbf{B}_a - \mathbf{L} \mathcal{B}^{-1} \mathbf{B}_a = \mathbf{0}.$$

The result above shows that $\mathbf{I}_n + \mathbf{V}^\top \mathcal{M}^{-1} \mathbf{U} = \mathbf{I}_n$ is regular and we can use the Sherman-Morrison-Woodbury formula to solve (3).

5. SOLVING $\mathcal{M}\mathbf{c} = \mathbf{q}$ WITH THE SHERMAN-MORRISON-WOODBURY FORMULA

With (5) the solution of (3) can now be expressed as

$$\begin{aligned}\mathbf{c} &= \mathbf{M}^{-1}\mathbf{q} = (\mathcal{M} + \mathcal{U})^{-1}\mathbf{q} = \mathcal{M}^{-1}\mathbf{q} - \mathcal{M}^{-1}\mathbf{U}(\mathbf{I}_n + \mathbf{V}^\top \mathcal{M}^{-1}\mathbf{U})^{-1}\mathbf{V}^\top \mathcal{M}^{-1}\mathbf{q} \\ &= \mathcal{M}^{-1}\mathbf{q} - \mathcal{M}^{-1}\mathbf{U}\mathbf{V}^\top \mathcal{M}^{-1}\mathbf{q} = \mathcal{M}^{-1}\mathbf{q} - \mathcal{M}^{-1}\mathcal{U}\mathcal{M}^{-1}\mathbf{q}.\end{aligned}$$

This gives us an algorithm to solve (3):

- (1) Solve $\mathcal{M}\boldsymbol{\xi} = \mathbf{q}$.
- (2) Solve $\mathcal{M}\boldsymbol{\zeta} = \mathcal{U}\boldsymbol{\xi}$.
- (3) Calculate $\mathbf{c} = \boldsymbol{\xi} - \boldsymbol{\zeta}$.

First we study the problem (1.) in detail. We have to solve

$$\begin{bmatrix} \mathbf{Y}_0 & \mathbf{I} \\ \mathbf{Y}_1 & \mathbf{I} \\ \ddots & \ddots \\ & \mathbf{Y}_{m-2} & \mathbf{I} \\ & & \mathcal{B} \end{bmatrix} \begin{bmatrix} \boldsymbol{\xi}_0 \\ \boldsymbol{\xi}_1 \\ \vdots \\ \boldsymbol{\xi}_{m-2} \\ \boldsymbol{\xi}_{m-1} \end{bmatrix} = \begin{bmatrix} \mathbf{q}_0 \\ \mathbf{q}_1 \\ \vdots \\ \mathbf{q}_{m-2} \\ \mathbf{q}_{m-1} \end{bmatrix}.$$

Therefore we solve $\mathcal{B}\boldsymbol{\xi}_{m-1} = \mathbf{q}_{m-1}$ and use recursion to find the other $\boldsymbol{\xi}_j$:

$$\mathbf{Y}_j\boldsymbol{\xi}_j = \mathbf{q}_j - \boldsymbol{\xi}_{j+1}, \quad j = m-2, \dots, 0.$$

We use the same method for our problem (2.). After we calculated

$$\mathcal{U}\boldsymbol{\xi} = \begin{bmatrix} \mathbf{0} & \cdots & \mathbf{0} \\ \ddots & & \\ \mathbf{B}_a & & -\mathbf{B}_a\mathbf{L} \end{bmatrix} \begin{bmatrix} \boldsymbol{\xi}_0 \\ \vdots \\ \boldsymbol{\xi}_{m-1} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{B}_a(\boldsymbol{\xi}_0 - \mathbf{L}\boldsymbol{\xi}_{m-1}) \end{bmatrix},$$

the resulting system of linear equations is

$$\begin{bmatrix} \mathbf{Y}_0 & \mathbf{I} & & \\ \mathbf{Y}_1 & \mathbf{I} & & \\ \ddots & \ddots & \ddots & \\ & \mathbf{Y}_{m-2} & \mathbf{I} & \\ & & \mathcal{B} & \end{bmatrix} \begin{bmatrix} \boldsymbol{\zeta}_0 \\ \boldsymbol{\zeta}_1 \\ \vdots \\ \boldsymbol{\zeta}_{m-2} \\ \boldsymbol{\zeta}_{m-1} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{B}_a(\boldsymbol{\xi}_0 - \mathbf{L}\boldsymbol{\xi}_{m-1}) \end{bmatrix}.$$

Again we first solve $\mathcal{B}\boldsymbol{\zeta}_{m-1} = \mathbf{B}_a(\boldsymbol{\xi}_0 - \mathbf{L}\boldsymbol{\xi}_{m-1})$ and then solve the remaining systems of linear equations with recursion:

$$\mathbf{Y}_j\boldsymbol{\zeta}_j = -\boldsymbol{\zeta}_{j+1}, \quad j = m-2, \dots, 0.$$

6. CONDENSING

We want to compare the new method above with the well known standard method from Stoer and Bulirsch. They solve (3) in the following way (see [1] or [4]):

- (1) Compute $\mathbf{E} := \mathbf{B}_a + \mathbf{B}_b\mathbf{X}_{m-1} \cdots \mathbf{X}_0$ and $\mathbf{u} := \mathbf{q}_{m-1} - \mathbf{B}_b\mathbf{X}_{m-1}(\mathbf{q}_{m-2} + \mathbf{X}_{m-2}\mathbf{q}_{m-3} + \cdots + \mathbf{X}_{m-2} \cdots \mathbf{X}_1\mathbf{q}_0)$.
- (2) Solve $\mathbf{E}\mathbf{c}_0 = \mathbf{u}$.
- (3) Compute the remaining \mathbf{c}_j with recursion: $\mathbf{c}_{j+1} = \mathbf{q}_j + \mathbf{X}_j\mathbf{c}_j$.

In the first step of our new algorithm from the section above we solve $\mathcal{B}\boldsymbol{\xi}_{m-1} = \mathbf{q}_{m-1}$. Notice that $\mathcal{B} = \mathbf{N}\mathbf{L}$. But $\mathbf{N} = \mathbf{E}$ holds. This follows directly from (4) and (7). That means our new algorithm has the same stability problems like the Condensing method. See [1] and [3] for a detailed discussion.

Therefore we only analyse the number of flops used by the two algorithms to compare them.

TABLE 1. Running time analysis for the Condensing method.

step	description	flops
1	Compute \mathbf{E} and \mathbf{u} . Because we compute the products of the \mathbf{X}_j matrices in \mathbf{E} we can use them to compute \mathbf{u} , too. Therefore we need no extra product computations of matrices to compute \mathbf{u} .	
	• $m - 1$ matrix-matrix multiplications for \mathbf{E}	$(m - 1)(2n^3 - n^2)$
	• one matrix addition for \mathbf{E}	n^2
	• $m - 1$ matrix-vector products for \mathbf{u}	$(m - 1)(2n^2 - n)$
	• m vector additions for \mathbf{u}	mn
2	Solve $\mathbf{E}\mathbf{c}_0 = \mathbf{u}$.	$2/3n^3$
3	Compute the remaining \mathbf{c}_j with recursion.	
	• $m - 1$ matrix-vector products	$(m - 1)(2n^2 - n)$
	• $m - 1$ vector additions	$(m - 1)n$
\sum	$= 2mn^3 + 3mn^2 - 4/3n^3 - 2n^2 + n$ flops	

TABLE 2. Running time analysis of our new method.

step	description	flops
1	Solve $\mathcal{M}\xi = \mathbf{q}$.	
1.1	Solve $\mathcal{B}\xi_{m-1} = \mathbf{q}_{m-1}$.	
	• Compute $\mathbf{T} := \mathbf{L}^{-1} = \mathbf{X}_{m-2} \cdots \mathbf{X}_0$.	$(m - 2)(2n^3 - n^2)$
	• Compute $\mathbf{N} := \mathbf{B}_a + \mathbf{B}_b \mathbf{X}_{m-1} \mathbf{T}$.	$4n^3 - n^2$
	• Solve $\mathbf{Ns} = \mathbf{q}_{m-1}$.	$2/3n^3$
	• Compute $\xi_{m-1} = \mathcal{B}^{-1} \mathbf{q}_{m-1} = \mathbf{L}^{-1} \mathbf{N}^{-1} \mathbf{q}_{m-1} = \mathbf{T}s$.	$2n^2 - n$
1.2	Use recursion to find the other ξ_j .	$(m - 2)(2/3n^3 + n)$
2	Solve $\mathcal{M}\zeta = \mathcal{U}\xi$.	
2.1	Solve $\mathcal{B}\zeta_{m-1} = \mathbf{B}_a(\xi_0 - \mathbf{L}\xi_{m-1})$.	
	• Solve $\mathbf{Tt} = \xi_{m-1}$.	$2/3n^3$
	• Compute $\mathbf{B}_a(\xi_0 - t)$.	$2n^2$
	• Solve $\mathbf{Ns} = \mathbf{B}_a(\xi_0 - t)$.	$2/3n^3$
	• Compute $\zeta_{m-1} = \mathbf{Ts}$.	$2n^2 - n$
2.2	Use recursion to find the other ζ_j .	$(m - 2)(2/3n^3)$
3	Compute $\mathbf{c} = \xi - \zeta$.	mn
\sum	$= 10/3mn^3 - mn^2 + mn - 2/3n^3 + 7n^2 - 4n$ flops	

7. RUNNING TIME ANALYSIS

We use LU-factorization to solve the systems of linear equations. We assume that this needs $2/3n^3$ flops for a $n \times n$ system.

The running time of the Condensing method is analyzed in Table 1. For a running time analysis of our new method see Table 2. The result is: The Condensing method is faster than the new method described above.

8. CONCLUSION

We found a new algorithm to solve the system of linear equations from the boundary and continuity conditions with the Sherman-Morrison-Woodbury formula. This new method has the same stability problems like the Condensing method. Our new method is also slower than the Condensing method. Therefore it is not recommendable to use the Sherman-Morrison-Woodbury formula in this case.

REFERENCES

- [1] P. DEUFLHARD AND F. BORNEMANN, *Numerische Mathematik 2, Gewöhnliche Differentialgleichungen*, de Gruyter, 2008.
- [2] M. HERMANN, *Numerik gewöhnlicher Differentialgleichungen*, Oldenbourg Verlag, 2004.

- [3] M. HERMANN, *Ein ALGOL-60-Programm zur Diagnose numerischer Instabilität bei Verfahren der linearen Algebra*, in Wiss. Ztschr. HAB Weimar, pp. 325–330 (1975).
- [4] J. STOER AND R. BULIRSCH, *Einführung in die Numerische Mathematik II*, Springer, 1973.

MATHEMATICAL INSTITUTE, UNIVERSITY OF JENA, D-07737 JENA, GERMANY.

E-mail address: Ivo.Hedtke@uni-jena.de